Studying the influence of regulatory mechanisms in the treatment of Myeloproliferative Neoplasms

Abstract. Myeloproliferative neoplasms (MPN) are clonal disorders of hematopoietic stem cells caused by driver mutations, essentially in JAK2. IFNα therapy has shown promising results in MPN blood cancers, leading in some cases to molecular remission. However, the mechanism of action of IFNα remains unclear. To better understand the effect of IFNα on MPN patients, our team previously proposed a mathematical model of hematopoiesis calibrated using longitudinal data from a cohort of patients. A potential limitation in our model was the absence of feedback regulation. In this article, we study several ways to implement regulatory mechanisms. Using Bayesian statistical inference and a model selection procedure, we find that our data are consistent with the hypothesis that the production of wild-type polynuclear neutrophils (PNN) might be regulated through red blood cells. However, we only slightly improve our previous results, suggesting that regulatory mechanisms would only play a minor role in the hematopoietic dynamics of MPN under IFNα. More data would be required to validate our hypothesis. If confirmed, our findings would imply that the measure of the variant allele frequency (VAF) among PNN, as done in clinical routine for patient monitoring, might not be an appropriate proxy of the VAF among progenitor cells, nor a good estimator of the instant effect of IFNα.

Keywords: Feedback Regulation · Bayesian Model Selection · Hematopoiesis · IFNα Therapy · Myeloproliferative Neoplasms

1 Introduction

Myeloproliferative Neoplasms (MPN) are malignant hematological pathologies resulting in the overproduction of matures blood cells and the deregulation of hematopoiesis. The disease is often detected belatedly after complications such as thrombosis or cardiovascular events and can degenerate to acute leukemia. These blood cancers occur following the acquisition of a specific somatic mutation in a hematopoietic stem cell (HSC). The primary driver mutation of the MPN disease affects the JAK2 protein (mutation $JAK2^{V617F}$) that plays a crucial role in cell signaling [2]. Following homologous recombination, homozygous malignant subclones can develop in parallel to the heterozygous ones.
Advances in the understanding of this disease are crucial to enable the development of treatments that will lead to a patient's recovery. Interferon alpha (IFNα), a natural inflammatory cytokine that has long been used to treat many diseases, has shown promising results in MPN. Indeed, IFNα induces not only

a hematological response, i.e., a normalization in blood cell counts, but also a molecular response, i.e., a reduction of mutated cells [4, 5].

In order to enable the development of personalized medicine, it is necessary to be able to understand and quantify the impact of this treatment on (mutated) hematopoietic cells. In previous work, Mosca et al. [1] tried to achieve this goal by building a mathematical model, calibrating it based on longitudinal observations from a cohort of patients, and finally inferring the HSC dynamics under IFNα. Their model proved to describe most of the patient's dynamics adequately, yet some limitations and subsequent avenues for improvement were identified. In particular, their model did not integrate any regulatory mechanisms that might be relevant to explain why some patients exhibit high measures of variant allele frequency (VAF) among polynuclear neutrophils (PNN). Feedback regulations have been widely used for modelling hematopoiesis, as by Marciniak et al. [6] or Jiao et al. [7]. In such models, some mature, fully differentiated hematopoietic cells influence their own production through a metabolic pathway involving cytokines.

In this article, we explore several ways to consider regulatory mechanisms and study their influence on the treatment of MPN. We derive different alternative models and use a Bayesian model selection procedure based on the Akaike information criterion (AIC) and the deviance information criterion (DIC). We apply this procedure using data from one illustrative patient. Then, we estimate the parameters of the selected model for several patients of the cohort.

2 Models

2.1 Based model without regulation

The model of Mosca et al. [1], without regulation, is the based model that we will further extend to include regulatory mechanisms. The description in this section was presented in more details in their previous work.

One brick for one cellular type. To describe cell differentiation and proliferation dynamics, Mosca et al. [1] proposed a compartmental model in which each compartment represents a category of cells in the hematopoietic hierarchical structure. The first assumption was the absence of interaction between wild-type (wt), heterozygous (het), and homozygous (hom) mutated cells. For each of these three genotypes, the authors ended up with a system of ordinary differential equations (ODE), as presented in eq. (1) and schematised in black in fig. 1, with parameter values specific to the cell type under consideration.

In this model, PNN die at the rate δ_{PNN}. Progenitors leave their compartment at the rate δ_i. Parameter κ_m models the proliferation of cells from progenitors to mature cells. HSCs can be quiescent (compartment 1) or active (compartment 2). In the latter case, HSCs can be recruited at a rate α to contribute to hematopoiesis. Once recruited, these cells divide. It was assumed that three division types occur at different rates (or probabilities). First, the HSC may
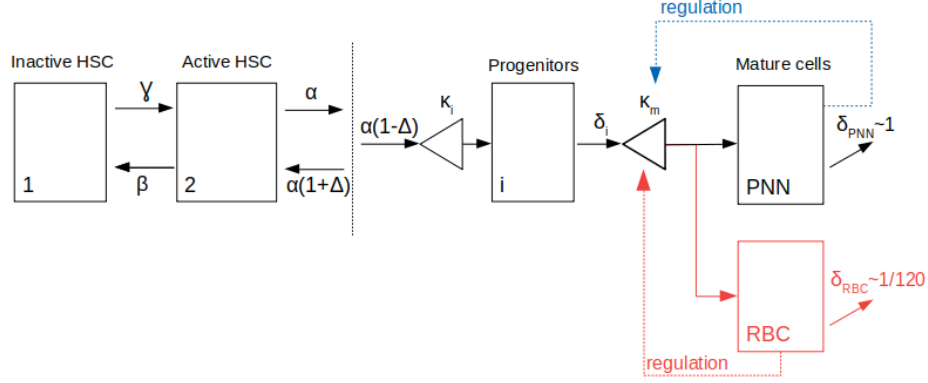
Fig. 1: Illustration of the cell compartments and parameters used in our different models (adapted from [1]). The four black compartments are those from the based model. We study the impact of regulatory mechanisms on the parameter κ_m that models the production of mature cells from progenitors. We study two hypotheses: either PNN are regulated through PNN (feedback regulation, blue line) or through red blood cells (red compartment) which have a much higher lifespan than PNN ($\sim$120 days vs $\sim$1 day).

divide asymmetrically (with probability p_1) and give rise to one HSC and one differentiated cell that further proliferates and becomes a progenitor. Second, the HSC may divide symmetrically (with probability p_2) and give rise to two HSC. Finally, the HSC may undergo a differentiated division and give rise to two differentiated cells (with probability $p_0 = 1 - p_2 - p_1$). Parameter $\Delta := p_2 - p_0$, which expresses the balance between these three division mechanisms, will be used in the following. Finally, the rates γ and β model exchanges between compartments 1 and 2. In the end, the system of equations to describe the dynamics of a given hematopoietic cell population (wt, het or hom) was as follows:

$$
\begin{cases}
\frac{dN_1(t)}{dt} & = -\gamma N_1(t) + \beta N_2(t) \\
\frac{dN_2(t)}{dt} & = \gamma N_1(t) + (\alpha \Delta - \beta) N_2(t) \\
\frac{dN_i(t)}{dt} & = \alpha(1 - \Delta)\kappa_i N_2(t) - \delta_i N_i(t) \\
\frac{dN_{PNN}(t)}{dt} & = \delta_i \kappa_m N_i(t) - \delta_{PNN} N_{PNN}(t)
\end{cases}
\tag{1}
$$

With $N_1(t)$, $N_2(t)$, $N_i(t)$ and $N_{PNN}(t)$ the numbers of cells, of a given genotype, in compartments 1, 2, immature and mature (PNN) respectively at time t. Appropriate indices should be used - both for parameters and numbers of cells - whether we refer to wt, het or hom cells.

Considering all cell types. The dynamics described above models the behaviour of a cell population of a given genotype. Yet, data from Mosca et al. [1] do not provide information about absolute values for quantities of wt, het or hom

cells separately, but rather their relative proportions. To take this into account, Mosca et al. considered that each population of a given genotype followed an ODE system as described in eq. (1). Then, they considered as outputs of their model no longer the numbers of cells but the proportions of immature heterozygous cells $z_{het}(t) = \frac{N_{i,het}(t)}{N_i(t)+N_{i,het}(t)+N_{i,hom}(t)}$ (and similarly for hom cells) as well as the mature VAF among PNN $y(t) = \frac{0.5 \cdot N_{PNN,het}(t)+N_{PNN,hom}(t)}{N_{PNN}(t)+N_{PNN,het}(t)+N_{PNN,hom}(t)}$. For convenience, $k_{i,het}$ was defined such that $\kappa_{i,het} = k_{i,het}\kappa_i$ and likewise $k_{i,hom}, k_{m,het}$ and $k_{m,hom}$.

Effect of IFNα and initial conditions. Following the idea of Michor et al. [8], Mosca et al. [1] considered that IFNα acts by modifying the values of some parameters in the model. Time $t = 0$ corresponds to the beginning of the treatment. Before that time, equations (1) are still valid, but it is assumed that the homeostatic conditions are satisfied, i.e. the system is in a quasi-stationary state. This is of course verified for wt cells as soon as $\Delta = 0$, which gives the following initial conditions: $N_1(0) = \frac{\beta}{\beta+\gamma}N_{HSC}$, $N_2(0) = \frac{\gamma}{\beta+\gamma}N_{HSC}$, $N_i(0) = \frac{\kappa_i \alpha}{\delta_i}N_2(0)$ and $N_{PNN}(0) = \frac{\kappa_m \delta_i}{\delta_{PNN}}N_i(0)$ with N_{HSC} the total wild-type HSC number considered constant. For mutated cells, Mosca et al. assumed that $\Delta_{het} \approx \Delta_{hom} \approx 0^+$. They also introduce $\eta_{het} = \frac{N_{1,het}(0)+N_{2,het}(0)}{N_{HSC}}$ and $\chi_{het} = \frac{N_{2,het}(0)}{N_{1,het}(0)+N_{2,het}(0)}$ for expressing the initial conditions for het and hom cells. From $t = 0$, patients are under treatment. IFNα is assumed to modify the values of some parameters, potentially in different ways depending on the cell type. In terms of notation, the superscript $*$ is added to the parameters impacted by the drug. From $t \geq 0$, eq. (1) remains valid with new parameters, and there is an equilibrium shift that induces a new dynamics.

2.2 Regulatory functions

The based model of Mosca et al. [1] was linear, and no interaction between genotypes was assumed. To extend their model, we consider a feedback regulation carried on by mature cells (see blue or red lines in fig. 1). Feedback regulation is relatively standard in biological systems and has, for example, been studied by Marciniak et al. [6] in hematopoiesis. Here, we assume that feedback regulation only impacts the production of mature cells (and not stem cells). Indeed, Mosca et al. obtained good fits for the dynamics of progenitor cells with their based model (as we can see for example on the left of fig. 3), which suggested that, if regulations mechanisms had a role to play, it would rather be at the level of immature cells than HSCs.

Biologically, $JAK2^{V617F}$ mutation affects the cell signaling [2] and makes the mutated cell less dependent on cytokines and growth factors: mutated cells might escape regulatory controls, especially homozygous compared to heterozygous cells. Thus, we choose to model a feedback regulation for wt and, to a lesser extent, for het progenitors, assuming that $\bar{\kappa}_{m,wt} = \bar{\kappa}_{m,wt}(N_{m,het}(t) + N_{m,hom}(t))$ and $\bar{\kappa}_{m,het} = \bar{\kappa}_{m,het}(N_{m,hom}(t))$ are not constants anymore but functions of a

quantity of mutated mature cells (more precision in the next paragraph). Then, the ODE systems we consider become non-linear and coupling between mutated and wild-type cells is introduced. Actually, regulation in biological systems is not directly carried on by cells but by cytokines. However, as done, for example, by Marciniak et al. [6], we consider that the number of cytokines might be directly proportional to the number of mature cells so that we can assume that these cells directly intervene in the regulatory functions.

To note that, biologically, it would be more accurate to consider a feedback regulation based on the whole number of mature cells (and not only the mutated ones). Our approach is an approximation that allows us to simplify the expression of the initial conditions and easily get a lower regulation in heterozygous cells than wt cells as expected biologically.

We study two different functions classically used for modelling feedback regulation, either a Hill function:

$$\bar{\kappa}_m(N) = \frac{\kappa_m}{1 + (\rho N)^n} \tag{2}$$

or a sigmoid:

$$\bar{\kappa}_m(N) = 2\kappa_m \left(1 - \frac{1}{1 + \exp(-\rho N)} \right) \tag{3}$$

where in both cases, $\bar{\kappa}_m(0) = \kappa_m$ and $\lim_{N \to +\infty} \bar{\kappa}_m(N) = 0$. ρ is the regulatory parameter to estimate. In the Hill function, we also introduce parameter n that we choose to set constant and equal to 2.

2.3   Regulation through different cell types

Regulation through PNN. Polynuclear neutrophils (PNN) are white blood cells with a life span of about one day. In the data from Mosca et al. [1], these are the cells used for measuring the VAF among mature cells. A first hypothesis would be that PNN regulate their own production through a feedback regulation. This is schematized with the blue line in fig. 1. In that case, the quantity N that appears in eq. (2) and (3) is $N_{PNN,hom}$ or $N_{PNN,het} + N_{PNN,hom}$ according to whether the regulation is acting on het or wet cells respectively.

The ODE system for homozygous cells will remains the same as for the based model (eq. (1)). For wt and het cells, the last line of system (1) is changed, introducing the dependence of the cell number $N_{PNN,hom}$ (and also $N_{PNN,het}$ for wt cells) in $\bar{\kappa}_m$.

Regulation through RBC. Red blood cells (RBC) have a much higher lifespan than PNN (about 120 days). Erythropoiesis - i.e., the production of RBC - is known to be regulated through the action of the erythropoietin (EPO). This cytokine acts on cells that possess the corresponding receptor (EpoR). Until recently, there was no biological evidence suggesting that PNN (and progenitor committed towards the granulocyte lineage) might be regulated through EPO

(or equivalently through the intermediary of RBC) since the EpoR was thought to be found essentially on erythroid cells. However, recent findings in the mouse demonstrated that EpoR could be expressed in other hematopoietic cells, including progenitor cells [3]. These findings might be consistent with the hypothesis that RBC might regulate the production of PNN (in addition to their own production).

To test this second hypothesis, we adapt the based model of Mosca et al. [1] by considering an additional compartment describing RBC, as depicted in red in fig. 1. Since we have no data about these cells, we assume a similar behavior to PNN, except for the lifespan $1/\delta_{RBC} \gg 1/\delta_{PNN}$. The ODE system (1) is modified accordingly with an additional equation and the introduction of the interaction within genotypes through the feedback regulatory function $\bar{\kappa}_m$. We get for wild-type cells (omitting the subscript wt for clarity):

$$\begin{cases} \frac{dN_1(t)}{dt} & = -\gamma N_1(t) + \beta N_2(t) \\ \frac{dN_2(t)}{dt} & = \gamma N_1(t) + (\alpha\Delta - \beta)N_2(t) \\ \frac{dN_i(t)}{dt} & = \alpha(1-\Delta)\kappa_i N_2(t) - \delta_i N_i(t) \\ \frac{dN_{PNN}(t)}{dt} & = \delta_i \bar{\kappa}_m \left(N_{mut,RBC}(t)\right) N_i(t) - \delta_{PNN} N_{PNN}(t) \\ \frac{dN_{RBC}(t)}{dt} & = \delta_i \bar{\kappa}_m \left(N_{mut,RBC}(t)\right) N_i(t) - \delta_{RBC} N_{RBC}(t) \end{cases} \qquad (4)$$

with $N_{mut,RBC}(t) = N_{RBC,het}(t) + N_{RBC,hom}(t)$.

3 Statistical inference and model selection

3.1 Data

The data we consider are from Mosca et al. [1]. Briefly, several MPN patients have been followed during five years over an IFNα therapy. Data were collected approximately every four months, both for progenitor cells and mature cells.

Concerning mature cells, VAF among PNN was measured by Taqman allelic discrimination qPCR. We model the uncertainty on the observation using a Gaussian noise. Concerning immature cells: based on the surface marker CD34$^+$, progenitor cells were isolated from the others cells, purified, and sorted at one progenitor per well in 96-well plates. Each progenitor gave a progeny of cells (colony) after 10 to 15 days of culture. The genotyping of each colony by Taqman allelic discrimination qPCR enabled Mosca et al. to know the genotype of each progenitor retrospectively, allowing them to know precisely how many wt, het and hom immature cells were in each subsample. At a given time t_i, for a given patient, these numbers are denoted $\hat{n}_i, \hat{n}_{i,het}$ and $\hat{n}_{i,hom}$ respectively. Because the subsamples are of reduced size (around 200 progenitors), the mutated immature cell proportions deduced are not fully representative of those from the entire progenitor population within the patient's body. This induces a sampling noise that Mosca et al. [1] modeled using a multinomial law.

In this article, we consider only $JAK2^{V617F}$ patients with VAF measures that go beyond 50%. Indeed, for patients with low VAF, studying the effect of regulatory

mechanisms might no be as pertinent. We first compare our different models based on patient #32 who is an illustrative patient of the study from Mosca et al. Then, we estimate the parameters of the model we would have selected for patients #3, 12, 18, and 20 from the cohort of Mosca et al. [1].

3.2 Parameter estimation

To estimate the parameters of each model we study, we consider a Bayesian framework. Let $\mathcal{M}$ be the model considered and $\boldsymbol{\theta}$ the random vector of the parameters to be estimated. Our goal is to estimate the posterior distribution of $\boldsymbol{\theta}$ given the data:

$$\mathbb{P}[\boldsymbol{\theta}|\mathcal{D}, \mathcal{M}] \propto \mathbb{P}[\mathcal{D}|\boldsymbol{\theta}, \mathcal{M}]\ \mathbb{P}[\boldsymbol{\theta}] \tag{5}$$

with $\mathcal{D} = \left\{\hat{n}_{i,het}, \hat{n}_{i,hom}, \hat{n}_{i,wt}, \hat{\mathrm{VAF}}_i\right\}_{i \in \mathrm{I}}$ the set of all measures (collected at different timepoint t_i for $i \in \mathrm{I}$) for the considered patient. For clarity, we omit to mention $\mathcal{M}$ in the following. The expression of the likelihood is as follows:

$$\mathbb{P}[\mathcal{D}|\boldsymbol{\theta}] \overset{(1)}{=} \prod_{i \in \mathrm{I}} \mathbb{P}[(\hat{n}'_i, \hat{n}''_i, \hat{n}_i, \hat{\mathrm{VAF}}_i)\,|\boldsymbol{\theta}] \overset{(2)}{=} \prod_{i \in \mathrm{I}} \mathbb{P}[(\hat{n}'_i, \hat{n}''_i, \hat{n}_i)\,|\boldsymbol{\theta}]\ \mathbb{P}[\hat{\mathrm{VAF}}_i|\boldsymbol{\theta}] \tag{6}$$

with (1) because conditionally on $\boldsymbol{\theta}$, the measures are independent, and (2) because the measures for immature and mature cells are obtained through two independent experiments. We do not have an analytical expression for the posterior. We approximate the distribution by sampling from it with a Markov Chain Monte Carlo (MCMC) algorithm, more specifically the Metropolis-Hastings algorithm. It relies on a proposal distribution that we choose as a multivariate normal distribution with zero mean and $\boldsymbol{\Sigma}$ as covariance matrix. Potentially, the parameter space dimension is large. In that case, the Metropolis-Hatings algorithm often proves inefficient, as it becomes complicated to define a covariance matrix that allows a good convergence of the algorithm. Adaptative algorithms have been proposed to circumvent this problem [10]. We proposed an alternative and simpler method which proved very efficient in practice. We start by learning the covariance matrix of the proposal distribution and by choosing a starting point of the Markov chain located at the maximum a posteriori using the CMA-ES algorithm [9].

We made the code for estimating the parameters of our different models publicly available on GitLab[1]. Several assumptions have been made by Mosca et al. to further decrease the number of parameters to estimate so that the based model could be identifiable given the data they had. These assumptions are presented in more detail in [1] and implemented in our code. A summary table of the parameters is presented in Figure 2. We end up with seven parameters to estimate for the based model, which are presented in Table 2.

---

[1] link hidden to respect the blind peer review process

	Before treatment and initial conditions		Under IFNα ($t \geq 0$)	
	Parameter	Value	Parameter	Value
WT	α	$1/30$	α^*	$= \alpha$
	Δ	0	Δ^*	$= \Delta$
	γ	$1/300$	γ^*	$= \gamma$
	β	$= \gamma(1-\chi)/\chi$	β^*	$= \beta$
	κ_i	NR	$k_i^* = \kappa_i^*/\kappa_i$	NR
	δ_i	$1/6$	δ_i^*	$= \delta_i$
	κ_m	NR	$k_m^* = \kappa_m^*/\kappa_m$	NR
	δ_m	1	δ_m^*	$= \delta_m$
	χ	0.1		
	N_{HSC}	NR		
Het	α_{het}	$= \alpha$	α^*_{het}	$= \alpha$
	Δ_{het}	0	$\boldsymbol{\Delta^*_{het}}$	**To estimate**
	γ_{het}	NR	$\boldsymbol{\gamma^*_{het}}$	**To estimate**
	β_{het}	NR	β^*_{het}	$= \beta$
	$k_{i,het} = \kappa_{i,het}/\kappa_i$	1	$k_{i,het}^* = \kappa_{i,het}^*/\kappa_{i,het}$	$= k_i^*$
	$\delta_{i,het}$	$= \delta_i$	$\delta_{i,het}^*$	$= \delta_i$
	$\boldsymbol{k_{m,het} = \kappa_{m,het}/\kappa_m}$	**To estimate**	$k_{m,het}^* = \kappa_{m,het}^*/\kappa_{m,het}$	$= k_m^*$
	$\delta_{m,het}$	$= \delta_m$	$\delta_{m,het}^*$	$= \delta_m$
	χ_{het}	$= \chi$		
	$\boldsymbol{\eta_{het}}$	**To estimate**		
Hom	α_{hom}	$= \alpha$	α^*_{hom}	$= \alpha$
	Δ_{hom}	0	$\boldsymbol{\Delta^*_{hom}}$	**To estimate**
	γ_{hom}	NR	$\boldsymbol{\gamma^*_{hom}}$	**To estimate**
	β_{hom}	NR	β^*_{hom}	$= \beta$
	$k_{i,hom} = \kappa_{i,hom}/\kappa_i$	1	$k_{i,hom}^* = \kappa_{i,hom}^*/\kappa_{i,hom}$	$= k_i^*$
	$\delta_{i,hom}$	$= \delta_i$	$\delta_{i,hom}^*$	$= \delta_i$
	$k_{m,hom} = \kappa_{m,hom}/\kappa_m$	$= k_{m,het}$	$k_{m,hom}^* = \kappa_{m,hom}^*/\kappa_{m,hom}$	$= k_m^*$
	$\delta_{m,hom}$	$= \delta_m$	$\delta_{m,hom}^*$	$= \delta_m$
	χ_{hom}	$= \chi$		
	$\boldsymbol{\eta_{hom}}$	**To estimate**		

Fig. 2: Summary table of the parameters used in the based model (from Mosca et al. [1]).

3.3   Model selection

Once having estimated the parameters of each model we study, for patient #32, we want to compare them and select the one that performs best. For that purpose, we use two different criteria. The first one is the Akaike Information Criteria (AIC):

$$AIC = 2k - 2\log\left(\mathbb{P}[\mathcal{D}|\boldsymbol{\theta}_{MLE}, \mathcal{M}]\right) \tag{7}$$

with $\boldsymbol{\theta}_{MLE}$ the maximum likelihood estimator and k the number of parameters to estimate. To note that, since we consider uniform prior distributions, $\boldsymbol{\theta}_{MLE}$ is also the maximum posterior estimator. It could be estimated both using the output of the CMA-ES algorithm or based on the generated MCMC.

The second criterion we use is the deviance information criterion (DIC) [11], broadly used in Bayesian model selection problems:

$$DIC = D\left(\mathbb{E}[\boldsymbol{\theta}|\mathcal{D}, \mathcal{M}]\right) + 2p_D \tag{8}$$

With the deviance defined by $D(\boldsymbol{\theta}) = -2\log\left(\mathbb{P}[\mathcal{D}|\boldsymbol{\theta}, \mathcal{M}]\right)$ and p_D the effective number of parameters defined, following Gelman et al. [12], by $p_D = 0.5\mathbb{V}[D(\boldsymbol{\theta})]$.

4   Results

4.1   Red blood cells could regulate the production of mature cell

Model	Regulatory function	AIC	DIC
Based model	NA	132.1	142.4
Regulation through PNN	Hill function	133.8	144.6
	Sigmoid	133.6	140.5
Regulation through RBC	Hill function	132.1	138
	Sigmoid	132.3	146

Table 1: Results of the model selection procedure applied on patient #32.

To find how regulation could potentially impact the hematopoietic dynamics during IFNα therapy, we apply our model selection procedure using the data of patient #32, who is an illustrative patient in the study of Mosca et al. [1], with many data points and an interesting dynamic. Indeed, we observe a typical bell curve for this patient, with a first increase of the proportion of mutated cells among progenitors and PNN just after the start of the therapy, and then a slow decrease. Moreover, patient #32 exhibits a high VAF among PNN, with a proportion of mutated mature cells that remain at a high level even after the CF of progenitor cells decreases. Such behavior was poorly fitted with the based model (fig. 3, left). Intuitively, such observation suggests a regulatory mechanism that would influence the proliferation of mature cells. Besides, since the cells we
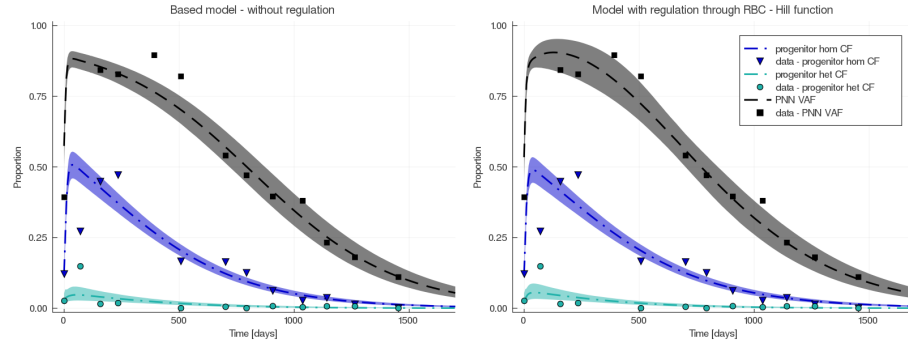
Fig. 3: Inferred dynamics for patient #32 from the based model (left) *vs* the model with a regulation through RBC and a Hill regulatory function (right). 95% credibility intervals are obtained by propagating uncertainties using a Monte-Carlo method.

are dealing with (progenitors and PNN) have a short lifespan of some days, it suggests a memory effect over several months that might be carried on by RBC (that have a higher lifespan of $\sim$120 days).

We tried to verify the plausibility of this hypothesis by applying our model selection procedure. We ran our Metropolis-Hasting algorithm (after using the CMA-ES algorithm) over 1 million iterations and a burn-in of 100,000. The results are presented in Table 1. Looking at the AIC, our alternative models with regulations perform all worse than the based model because of the penalization of the additional parameter to estimate. Only the model with feedback regulation through RBC and a Hill regulatory function performs as well as the based model. We get slightly different results by considering the DIC, with two models performing better than the based model and the model performing the best being the same as for the AIC.

Thus, our data are consistent with the hypothesis that the production of PNN would be regulated through RBC, the model that encapsulates this hypothesis getting the best performance. However, according to our criteria, we obtain only a slight improvement compared to the based model. Comparing the inferred dynamics between the based model and the best model in Figure 3, we still observe that the inferred dynamic for PNN looks visually better with the model implementing regulatory mechanisms. In particular, the 95% credibility interval reaches the two measures at 400 and 500 days, which was not the case for the based model. It might suggest that regulation plays a minor role and that a model implementing such effects might only be relevant if we had access to more observations.
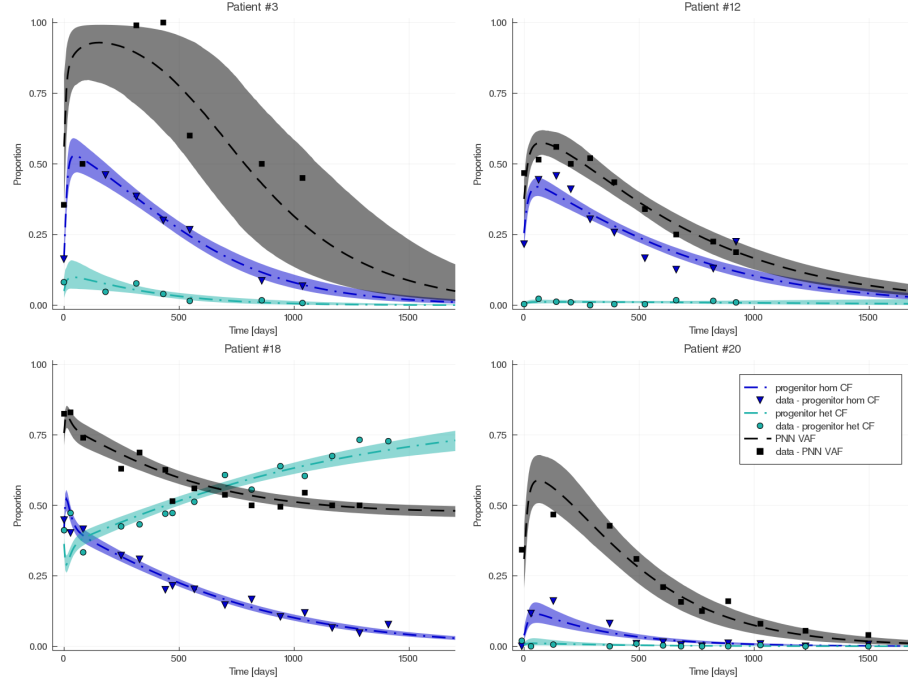
Fig. 4: Inferred dynamics for four additional patients (for the best model with feedback regulation through RBC). 95% credibility intervals are obtained by propagating uncertainties using a Monte-Carlo method.
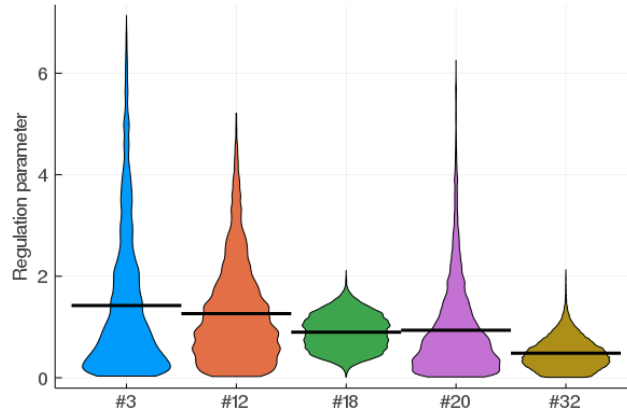

Fig. 5: Posterior distribution of the regulatory parameter ρ for each five patients we consider.

	#3	#12	#18	#20
$k_{m,hom}$	5.40 [1.61, 14.7]	1.61 [1.11, 2.02]	1.40 [1.01, 2.24]	10.46 [6.6, 15.1]
η_{hom}	0.23 [0.17, 0.31]	0.35 [0.26, 0.44]	1.87 [1.62, 2.1]	0.04 [0.02, 0.07]
γ^*_{hom}	0.06 [0.03, 0.1]	0.009 [0.006, 0.013]	0.004 [0.003, 0.004]	0.02 [0.01, 0.04]
Δ^*_{hom}	-0.14 [-0.2, -0.11]	-0.33 [-0.46, -0.23]	-0.86 [-0.99, -0.67]	-0.35 [-0.63, -0.17]
η_{het}	0.07 [0.03, 0.13]	0.01 [0.0, 0.02]	1.63 [1.41, 1.86]	0.01 [0.0, 0.01]
γ^*_{het}	0.03 [0.01, 0.08]	0.03 [0.0, 0.09]	0.0034 [0.003, 0.004]	0.02 [0.0, 0.06]
Δ^*_{het}	-0.40 [-0.95, -0.15]	-0.11 [-0.44, 0.02]	0.09 [0.05, 0.13]	-0.46 [-0.96, -0.08]
ρ	1.42 [0.03, 5.06]	1.26 [0.05, 3.77]	0.90 [0.3, 1.55]	0.94 [0.03, 3.45]

Table 2: Estimating the parameters of the best model for four additional patients. Expected values along with their 95% credibility interval are displayed.

4.2 Regulatory mechanisms would play a minor role in IFNα therapy

We then estimated the posterior distributions of the parameters (for the best model, i.e., with feedback regulation through RBC and a Hill regulatory function) for four other patients who exhibit a high VAF among PNN. The expected parameter values and the corresponding 95% credibility interval are presented in Table 2. In this article, we are mainly interested in the regulatory parameter ρ (see eq. (2) and (3)), which posterior distribution is displayed for each patient we considered in Figure 5. Regulatory mechanisms might play a more critical role for these four additional patients than patient #32. Indeed, we estimate higher mean values for parameter ρ. However, except for patient #18, values near 0 are reached with a high probability, confirming the low influence of regulatory mechanisms in MPN dynamics under IFNα. Comparing the AIC between the based model and the model with feedback regulation through RBC, we found that this latter model performed only better for patient #18.
The inferred dynamics for these four patients are presented in fig. 4. Compared to the results obtained by Mosca et al. [1], we get few differences in the inferred dynamics, except for patient #3. For this latter, the based model could not reach VAF about 100%, contrary to what we can observe with our model with regulation.

5 Discussion

Using mathematical modelling and Bayesian model selection, we explored how regulatory mechanisms could influence the dynamics of hematopoietic cells for MPN patients over IFNα therapy. We extended the work of Mosca et al. [1] by implementing different mechanisms that would regulate the production of PNN. We tested two main hypotheses: a feedback regulation carried on by PNN or regulation through RBC. The first hypothesis is a natural extension of the based model when the second one is based on recent findings having evidenced the existence of the EPO receptor on non-erythroid cells [3]. Our results were

mixed. Based on both the AIC and the DIC criteria, we found that the best model would be the one with regulation through RBC and a Hill regulatory function. However, we only slightly improved the results that we got with the based model, which does not allow us to conclude with confidence that RBC would truly regulate PNN. More reasonably, our results would suggest that regulation might play a role for some patients with high VAF among mature cells, but only a minor one. Given the observations from Mosca et al., our results suggest that it might not be relevant to implement regulatory mechanisms in their model since such implementation results in an additional parameter to estimate and does not strongly improve the data's fit. We also only explored a limited number of regulatory mechanisms; we can not exclude that another model would perform much better than ours. Biological regulatory mechanisms involve complex metabolic pathways, and appropriate modeling would probably require many parameters to estimate.

Finally, more data would be required to validate the hypothesis that RBC regulate the production of PNN during IFNα therapy. Such findings might have clinical implications. Since PNN have a short life span of about one day, their measure is used as a proxy of the instant effect of IFNα among progenitor and mature cells. However, regulated production of PNN through cells with a higher life span - in this case RBC - would imply that the VAF among PNN also embeds some information about RBC production months ago. Thus, VAF measures of PNN should be considered with caution.

References

1. Mosca, Blood (in review), 2021
2. Vainchenker, W., Kralovics, R. (2017). Genetic basis and molecular pathophysiology of classical myeloproliferative neoplasms. Blood, The Journal of the American Society of Hematology, 129(6), 667-679.
3. Zhang, H., Wang, S., Liu, D., Gao, C., Han, Y., Guo, X., ... & An, X. (2021). EpoR-tdTomato-Cre mice enable identification of EpoR expression in subsets of tissue macrophages and hematopoietic cells. Blood.
4. Yacoub, A., Mascarenhas, J., Kosiorek, H., Prchal, J. T., Berenzon, D., ... & Finazzi, M. C. (2019). Pegylated interferon alfa-2a for polycythemia vera or essential thrombocythemia resistant or intolerant to hydroxyurea. blood, 134(18), 1498-1509.
5. Gisslinger, H., Klade, C., Georgiev, P., Krochmalczyk, D., ... & Sivcheva, L. (2020). Ropeginterferon alfa-2b versus standard therapy for polycythaemia vera (PROUD-PV and CONTINUATION-PV): a randomised, non-inferiority, phase 3 trial and its extension study. The Lancet Haematology, 7(3), e196-e208.
6. Marciniak-Czochra, A., Stiehl, T., Ho, A. D., Jäger, W., Wagner, W. (2009). Modeling of asymmetric cell division in hematopoietic stem cells—regulation of self-renewal is essential for efficient repopulation. Stem cells and development, 18(3), 377-386.
7. Jiao, J., Luo, M., & Wang, R. (2018). Feedback regulation in a stem cell model with acute myeloid leukaemia. BMC systems biology, 12(4), 75-83.
8. Michor, F., Hughes, T. P., Iwasa, Y., Branford, S., Shah, N. P., Sawyers, C. L., Nowak, M. A. (2005). Dynamics of chronic myeloid leukaemia. Nature, 435(7046), 1267-1270.

9. Hansen, N. (2006). The CMA evolution strategy: a comparing review. In Towards a new evolutionary computation (pp. 75-102). Springer, Berlin, Heidelberg.

10. Andrieu, C., & Thoms, J. (2008). A tutorial on adaptive MCMC. Statistics and computing, 18(4), 343-373.

11. Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. Journal of the royal statistical society: Series b (statistical methodology), 64(4), 583-639.

12. Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). Bayesian Data Analysis Chapman & Hall. CRC Texts in Statistical Science.